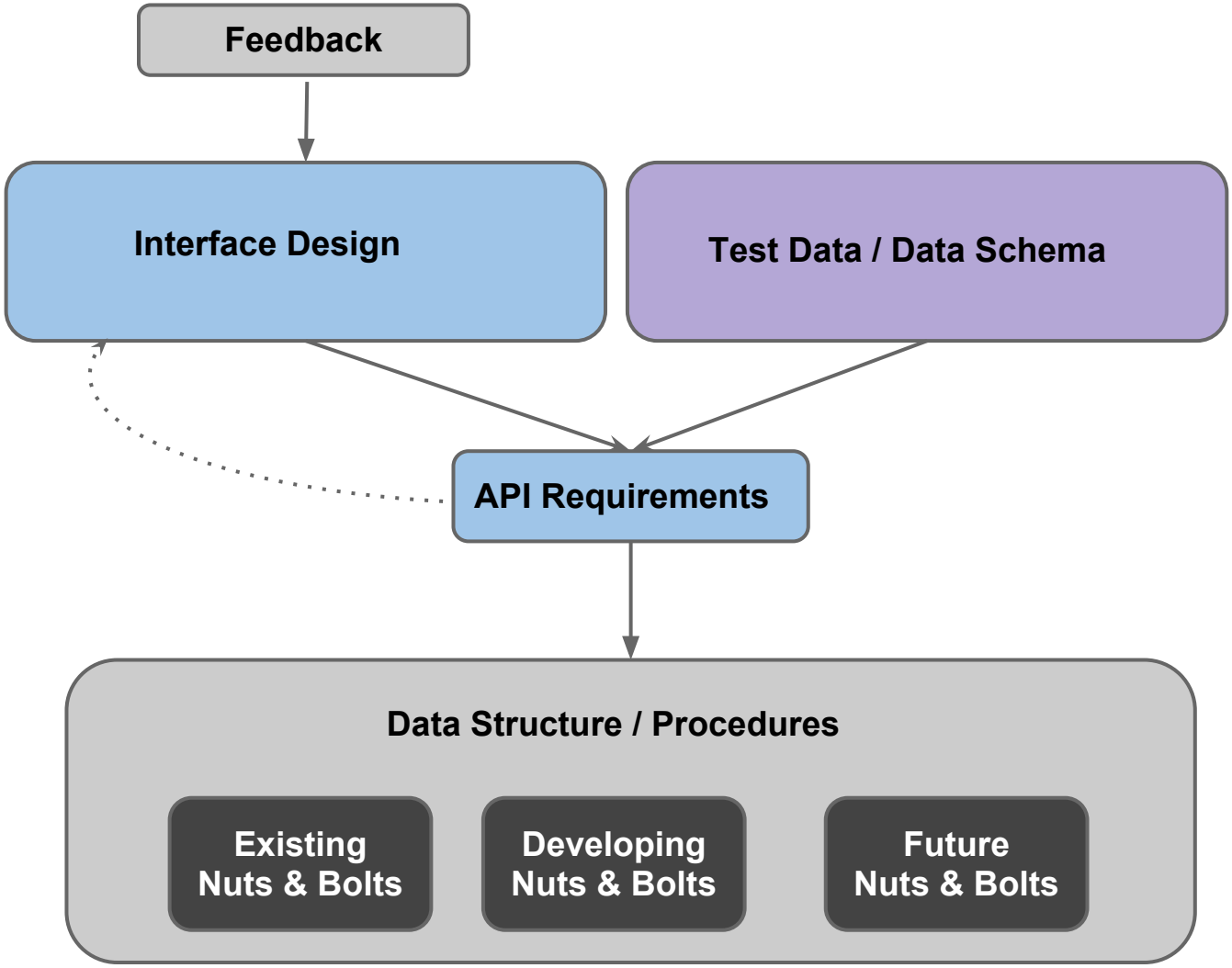
GINAS Software Development Update

- Backend:
  - API status
    - Entity persistence layer almost complete
    - Initial implementation of access control and security
    - Improved chemical cartridge based on evaluation of SRS test data
  - Loaded MySQL instance with SRS test data
    - Limited test of data migration with the API by next meeting
  - Prioritize development path in sync with the front end going forward

GINAS Software Development Update

- Frontend:
  - Registration mock-ups for Chemicals / Proteins
  - Defining and implementing necessary "widgets"
    - Reusable web components
  - Basic web interface in active development
    - Link to active development site very soon!
  - Driving force of the API development
  - Added new member Tongan Zhao to the team
    - Specialized in web-based technologies
    - Responsible for client side implementation

Development Strategy



Development Strategy

Focus on functionality (Top-Down)

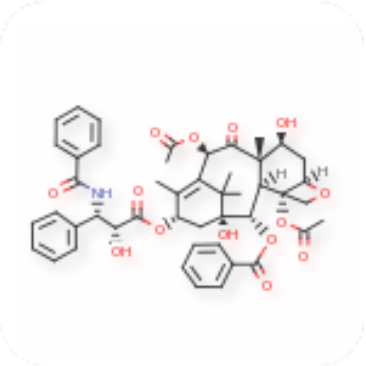
- What do users see?
- What do users *need* to do?
- What do users *want* to do?

Interface Design as discussion platform

- UI features determine API features
- API features determine Application features
- Application features guide model implementation

Case 1

Registering Small Molecule



Paclitaxel

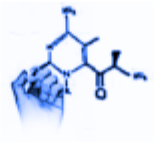
GINAS

Login Peryeat Password *********

Home Search Download Register Admin

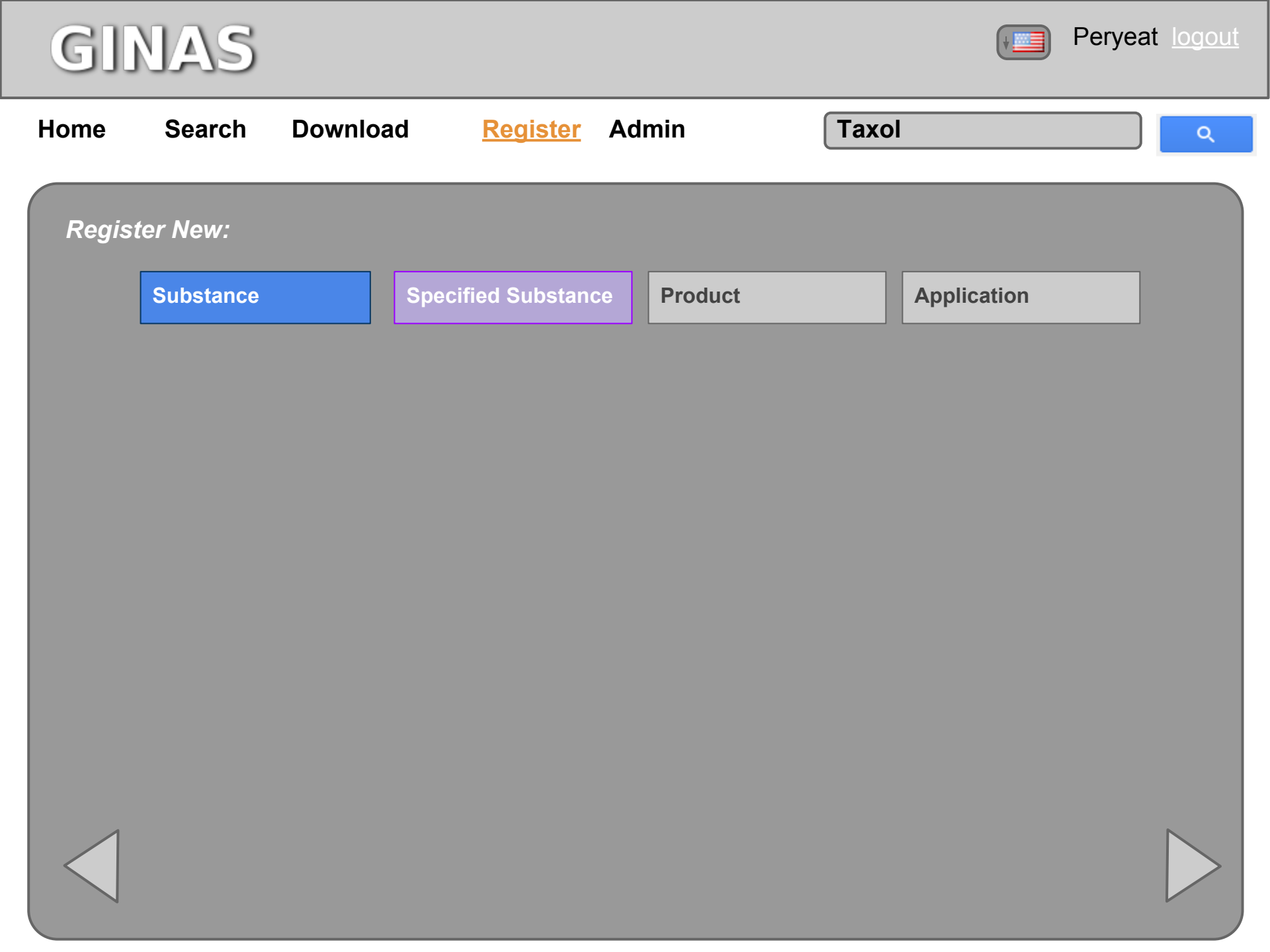
Search GINAS Taxol

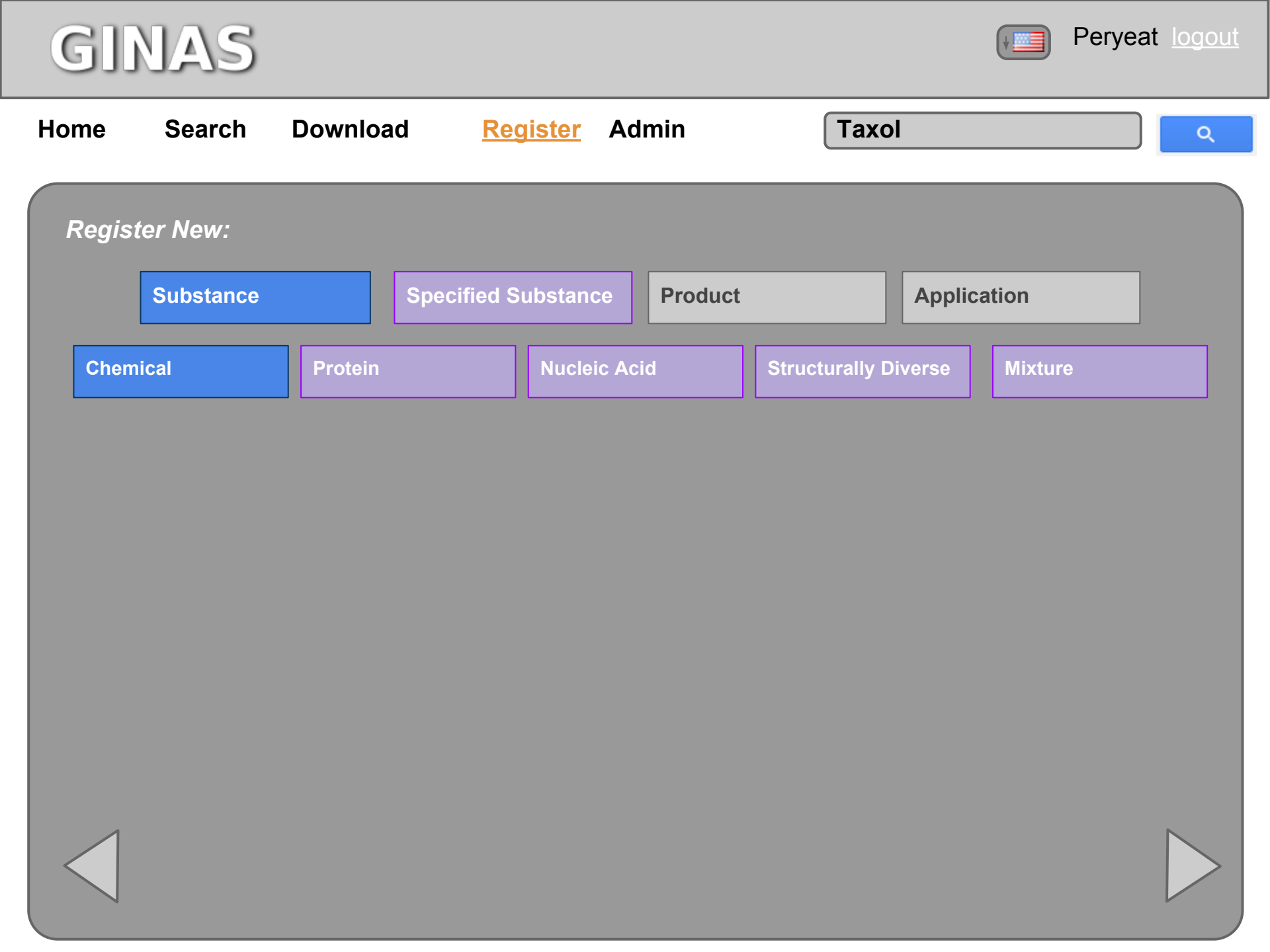
Chemical Structure

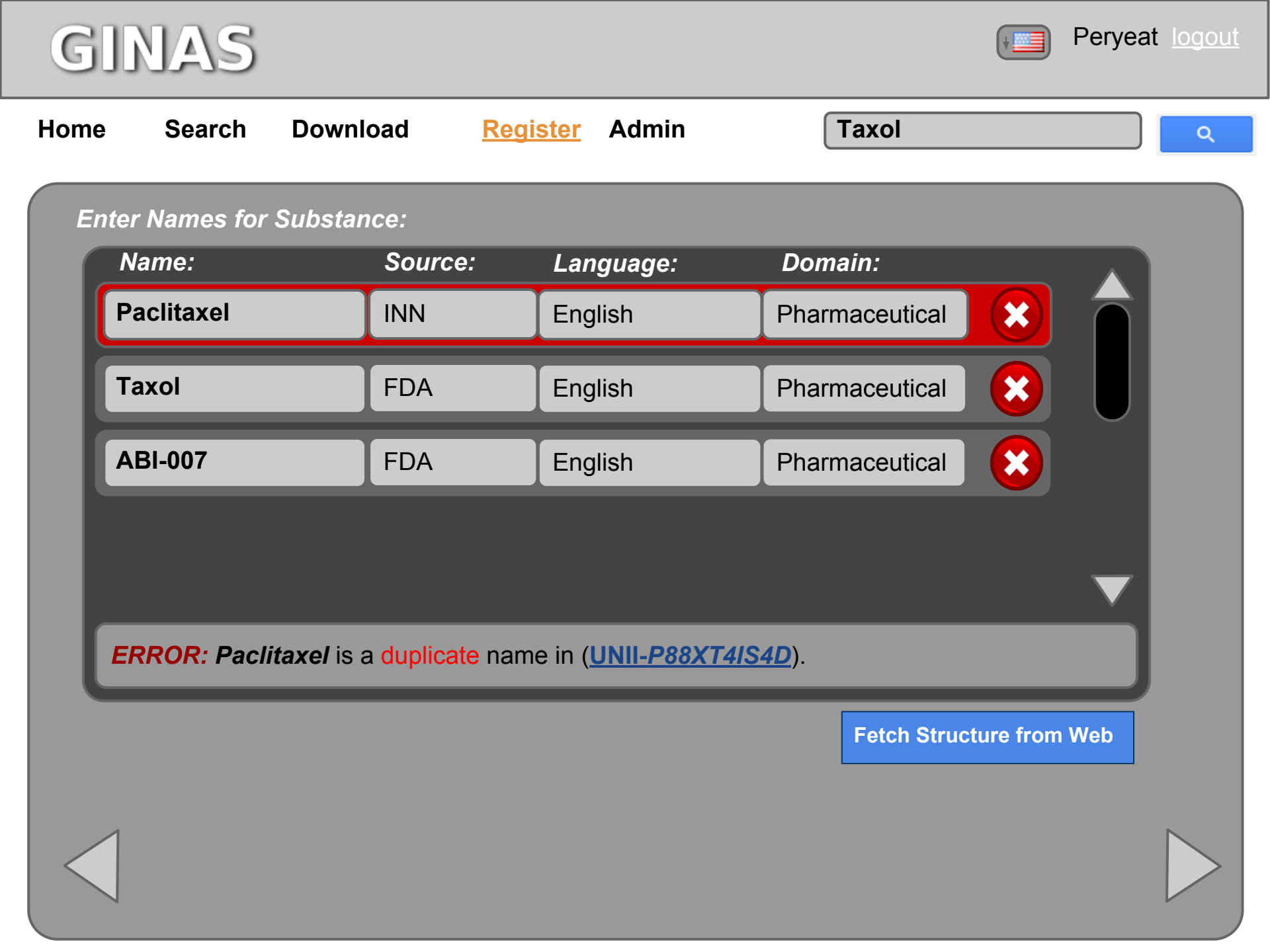


Amino Acid Sequence

Nucleic Acid Sequence


GINAS
Peryeat logout
Home
Search
Download
Register
Admin
Taxol
Register New:
Substance
Specified Substance
Product
Application


GINAS
Peryeat logout
Home
Search
Download
Register
Admin
Taxol
Register New:
Substance
Specified Substance
Product
Application
Chemical
Protein
Nucleic Acid
Structurally Diverse
Mixture


GINAS
Peryeat logout
Home
Search
Download
Register
Admin
Taxol
Enter Names for Substance:
Name:
Source:
Language:
Domain:
Paclitaxel
INN
English
Pharmaceutical
Taxol
FDA
English
Pharmaceutical
ABI-007
FDA
English
Pharmaceutical
ERROR: Paclitaxel is a duplicate name in (UNII-P88XT4IS4D).
Fetch Structure from Web

GINAS

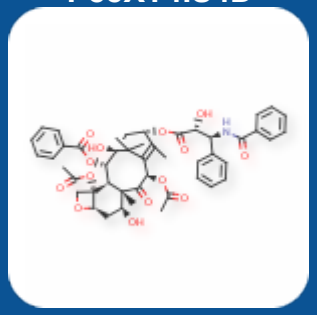
Peryeat logout

Home Search Download Register Admin Taxol

Enter Names for Substance:

Name:	*Source:*	*Language:*	*Domain:*	
Paclitaxel	INN	English	Pharmaceutical	✖
Taxol	FDA	Eng	rmaceutical	✖
ABI-007	FDA	Eng	rmaceutical	✖

P88XT4IS4D



Paclitaxel

ERROR: ***Paclitaxel*** is a duplicate name in (*UNII-P88XT4IS4D*).

Fetch Structure from Web

GINAS

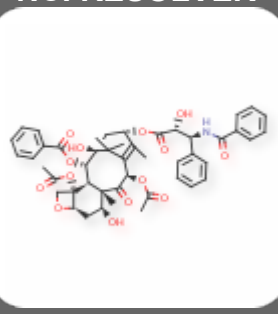
Peryeat logout

Home Search Download Register Admin Taxol

Found from other sources:

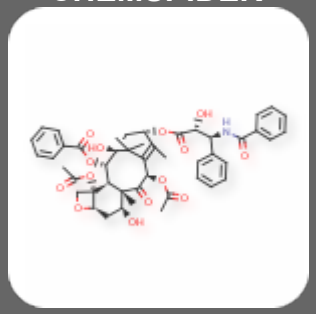
"Paclitaxel" (3) | "ABI-007" (2) | "Taxol" (5) | From Existing

NCI RESOLVER



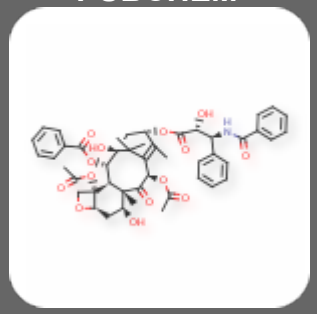
Choose

CHEMSPIDER



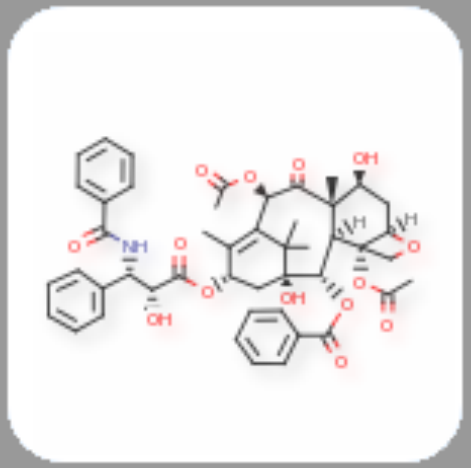
Choose

PUBCHEM



Choose

Skip



Defined Stereocenters: 11
Total Stereocenters: 11
E / Z centers: 0
Charge: 0
Molecular Formula:
$C_{47}H_{51}NO_{14}$
Molecular Weight:
853.906

Stereochemistry

Achiral | Absolute | Mixed | Racemic | Epimeric | Unknown

Optical Activity

(+ / -) | (+) | (-) | None | Unknown

Is there additional stereochemistry? Yes No

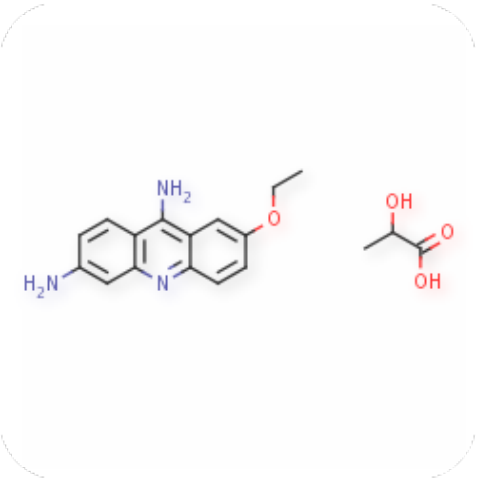
Register

Case 1 Requirements:

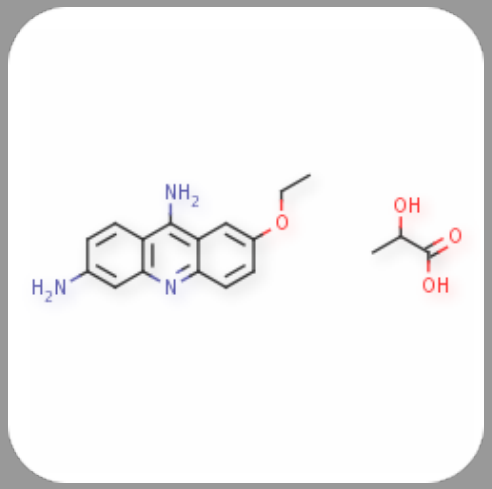
- Required API Features:
  - Given Name, return suitability for registration
    - OK / Warning / Error
    - Return referenced substances
  - Given Entity ID, return representation
  - Given Name, return representation from other databases
  - Given chemical structure, return breakdown:
    - Molecular Weight
    - Molecular Formula
    - Stereocenters defined / present
    - E/Z double bonds present

Case 2

Registering Small Molecule : Multiple Moieties

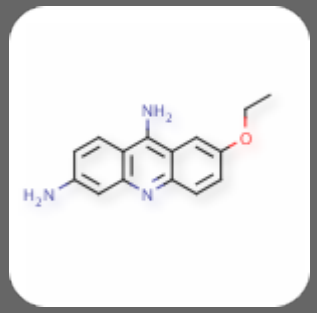


ETHACRIDINE LACTATE



There are 2 moieties for this compound:

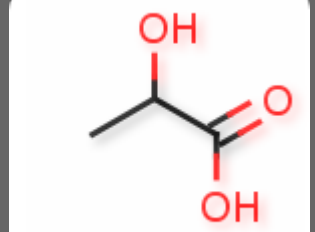
A19SF401CG



Moiety 1

x 1

UNREGISTERED



Moiety 2

x 1

Is the stoichiometry fixed as shown? Fixed Variable

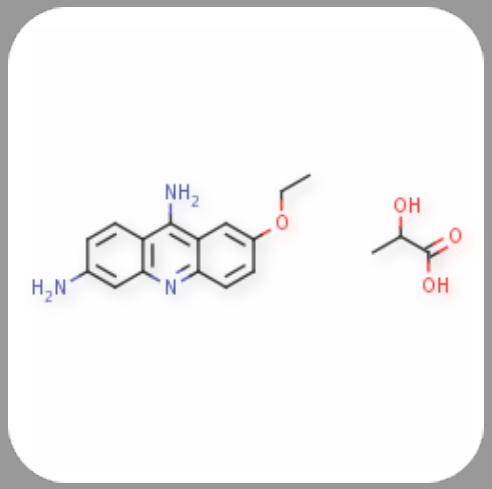
Defined Stereocenters: 0
Total Stereocenters: 1
E / Z centers: 0
Charge: 0

Molecular Formula:

$C_{15}H_{15}N_3O \cdot C_3H_6O_3$

Molecular Weight:

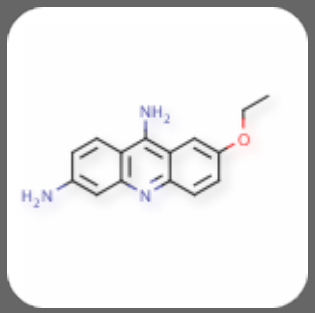
253.2991 • 90.08



Defined Stereocenters: 11
Total Stereocenters: 11
E / Z centers: 0
Charge: 0
Molecular Formula:
$C_{15}H_{15}N_3O \cdot C_3H_6O_3$
Molecular Weight:
253.2991 • 90.08

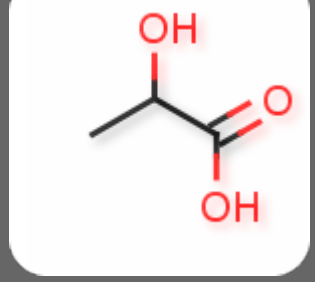
There are 2 moieties for this compound:

A19SF401CG



Moiety 1
x 1

UNREGISTERED



Moiety 2
x [1 - 4]

Is the stoichiometry fixed as shown? Fixed Variable

Normalize

	Avg:	Low:	High:
Moiety 1 :	1	1	1
Moiety 2 :	2.5	1	4

Case 2 Requirements:

- Required API Features:
  - Given chemical structure, return child moieties:
    - Named and ordered by decreasing molecular weight
    - With references to existing moieties, if they exist

Case 3

Registering Protein : Overview

GINAS

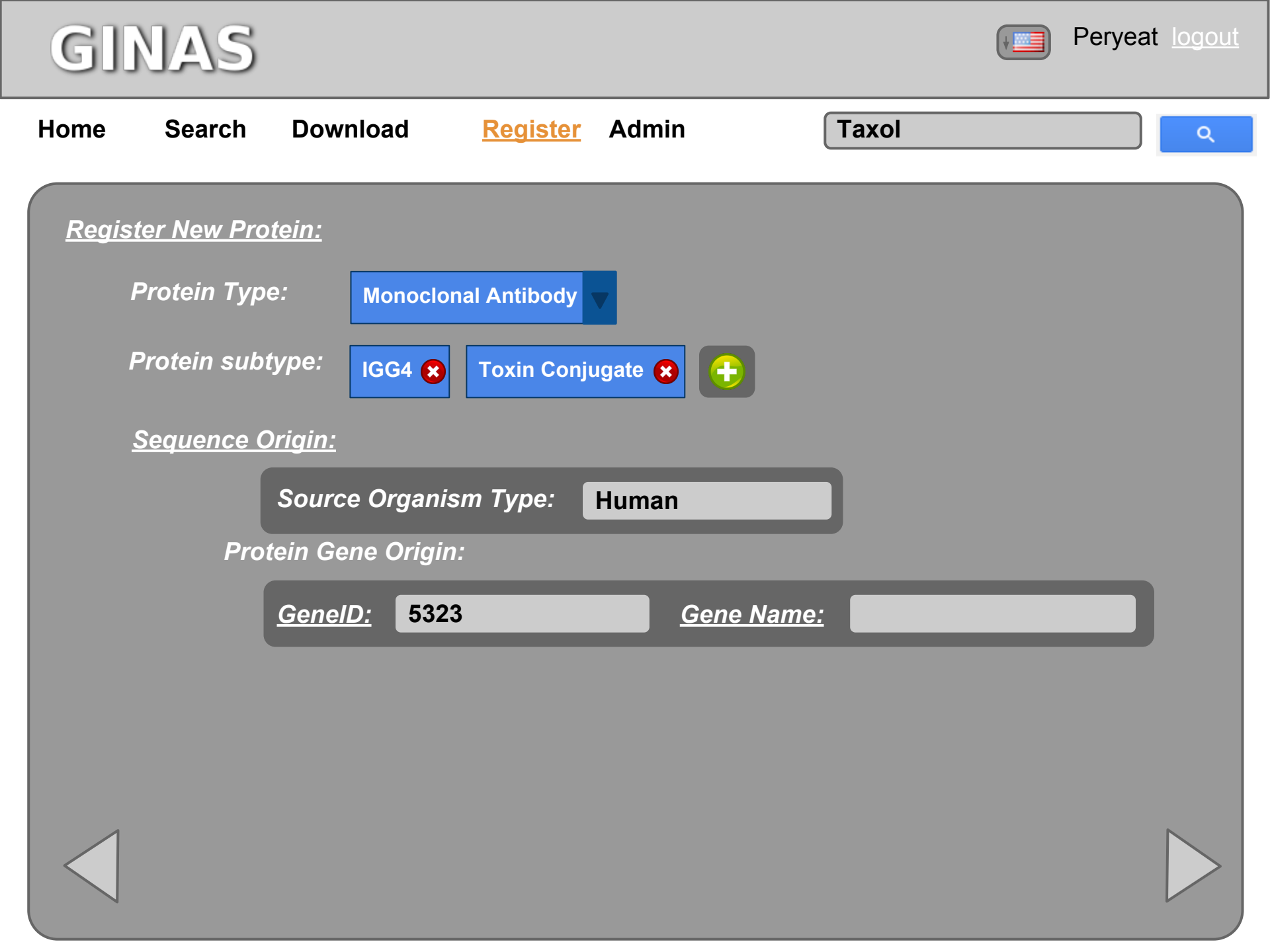
Peryeat logout

Home Search Download Register Admin Taxol

Register New:

Substance | Specified Substance | Product | Application

Chemical | Protein | Nucleic Acid | Structurally Diverse | Mixture


GINAS
Peryeat logout
Home
Search
Download
Register
Admin
Taxol
Register New Protein:
Protein Type:
Monoclonal Antibody
Protein subtype:
IGG4
Toxin Conjugate
Sequence Origin:
Source Organism Type:
Human
Protein Gene Origin:
GeneID:
5323
Gene Name:

Amino Acid Sequence:

Total number of Subunits: 2

Subunit 1: *Length: 80*

```
ASDASOWFNF ASDASOWFNF ACDASOWFNF ASDASOwfnF ASDASOCFNF
ASDASOWFCF ASDASOWFNF AXDASOWFNF
```

(D-Amino Acids) (X : Unnatural Amino Acid)

Subunit 2: *Length: 80*

```
ASDASOWFNF ASDASOWFNF ACDASOWFNF ASDASOWFNF ASDASOCFNF
ASDASOWFCF ASDASOWFNF ASDASOWFNF
```

Add Disulfide Links | Add Glycosylation | Add other Crosslink

Modifications:

Is there a modification to the substance?

Yes No

Modifications:

Is there a modification to the substance?

Yes No

Modification Type:

Structural Agent Physical

Modifications:

Is there a modification to the substance?

Yes | No

Modification Type:

Structural | Agent | Physical

Modification Location:

Site Specific | Residue Specific | Functional Group | C-Terminus | N-Terminus

Modifications:

Is there a modification to the substance?

Yes | No

Modification Type:

Structural | Agent | Physical

Modification Location:

Site Specific | Residue Specific | Functional Group | C-Terminus | N-Terminus

Acetylated | Pyro-glu | Myristoylated

Formylated | Palmitoylated | Other

Modifications:

Is there a modification to the substance?

Yes | No

Modification Type:

Structural | Agent | Physical

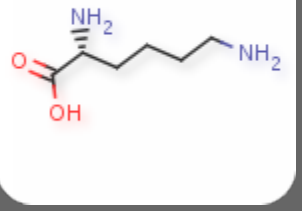
Modification Location:

Site Specific | Residue Specific | Functional Group | C-Terminus | N-Terminus

Amino Acid:	Lysine	
Average Number :	2.5	*/ 10*
High Number :	3.0	*/ 10*
Low Number :	2.0	*/ 10*

Define Modification

Lysine



Z43HSD024G

UNII: *Z43HSD024G*

Substitute

(must already exist)

Choose

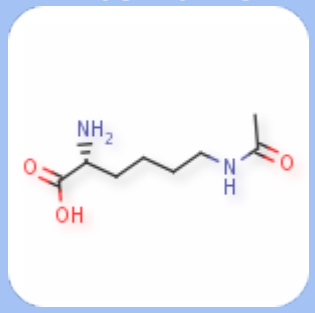
UNII:

Chose Amino Acid Substitution

Referenced by **Lysine:**

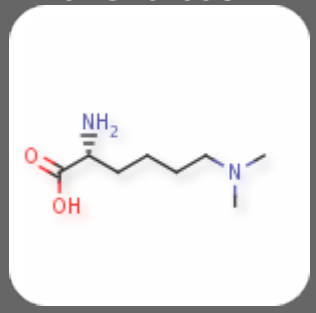
Search

A19SF401CG



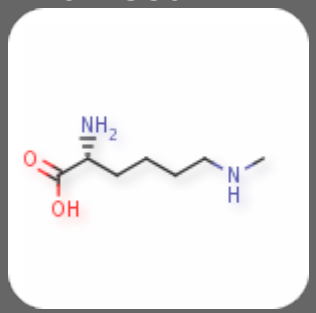
Acetyllysine

3YGF0495O2



(6-N,6-N)dimeth...

JD2SO34P21

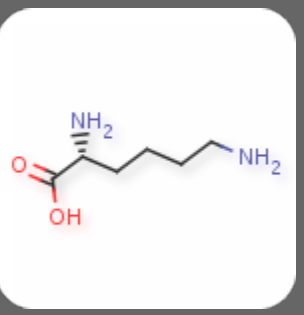


6-N-methyllysine

Done

Define Modification

Lysine



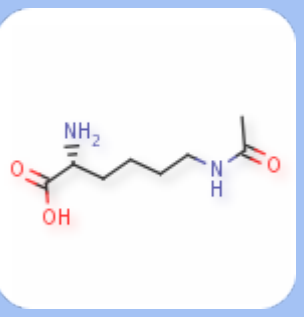
Z43HSD024G

UNII: *Z43HSD024G*

Substitute

(must already exist)

Acetyllysine



A19SF401CG

UNII: **A19SF401CG**

Case 3 Requirements:

- Required API Features:
  - Get all entities related to given entity.
    - With certain type of relationship
      - (e.g. Amino Acid Substitution)
  - Description of query / filter:
    - "Substances referenced by Lysine"

Case 3 Requirements:

- Required "web widgets":
  - Popup Substance searcher for quickly referencing other entities
  - Popup crosslinking tool for quickly defining disulfide bridges / other crosslinking agents